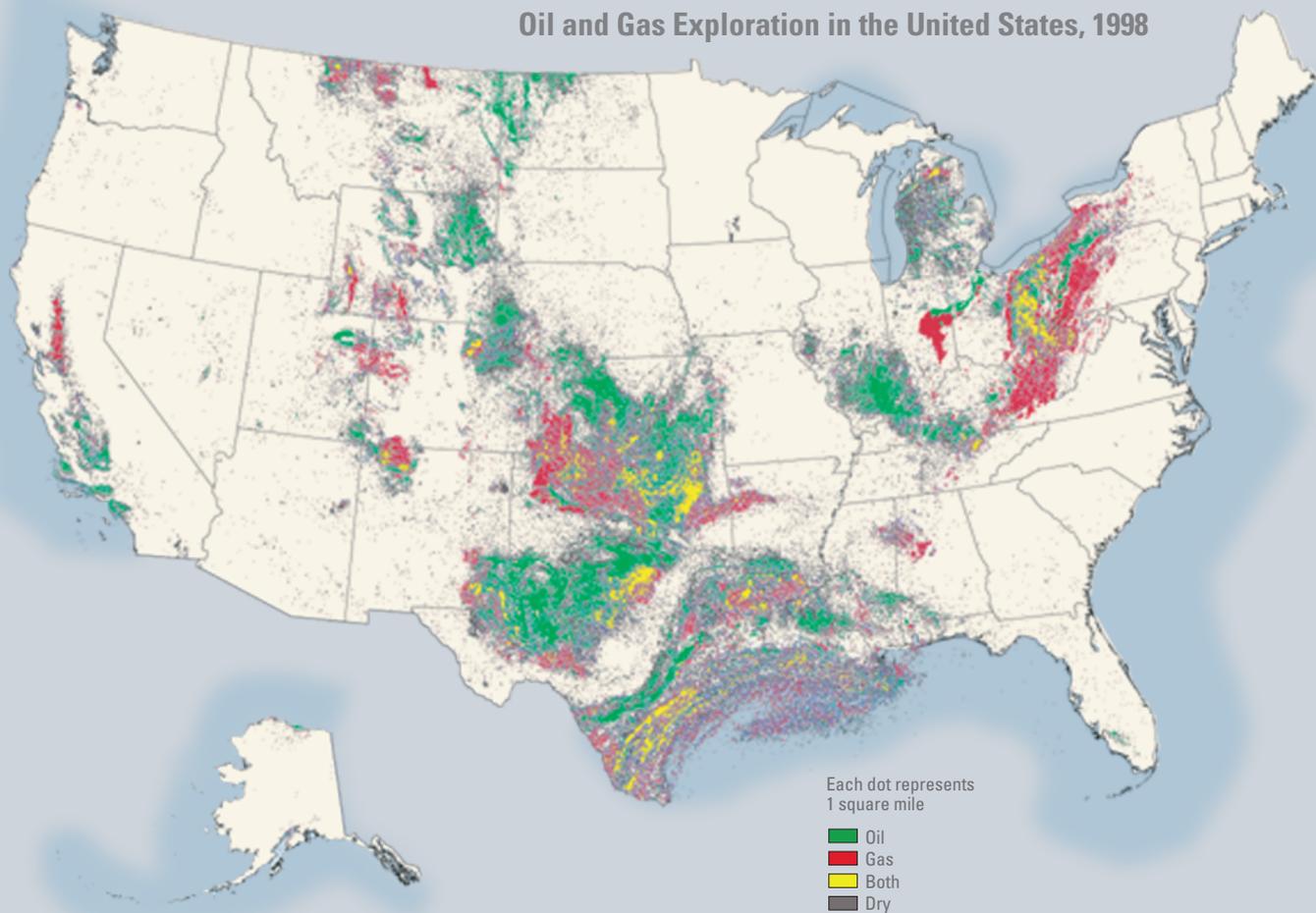


Uncertainty and Inferred Reserve Estimates— The 1995 National Assessment

Chapter G of
Geologic, Engineering, and Assessment Studies of Reserve Growth

U.S. Geological Survey Bulletin 2172–G



Cover. This map represents historical oil and gas exploration and production data for the conterminous United States and Alaska. It was derived from data used in U.S. Geological Survey Geologic Investigations Series I-2582.* The map was compiled using Petroleum Information Corporation's (currently IHS Corporation) database of more than 2.2 million wells drilled in the U.S. as of June 1993. The area of the U.S. was subdivided into 1 mi² grid cells for which oil and gas well completion data were available. Each colored symbol represents a 1 mi² cell (to scale) for which exploration has occurred. Each cell is identified by color as follows: red, a gas-producing cell; green, an oil-producing cell; yellow, an oil- and gas-producing cell; gray, a cell that has been explored through drilling, but no production has been reported. Mast and others (1998) gives details on map construction.

*Mast, R.F., Root, D.H., Williams, L.P., Beeman, W.R., and Barnett, D.L., 1998, Areas of historical oil and gas exploration and production in the conterminous United States: U.S. Geological Survey Geologic Investigations Series I-2582, one sheet.

Uncertainty and Inferred Reserve Estimates— The 1995 National Assessment

By Emil D. Attanasi and Timothy C. Coburn

U.S. Geological Survey Bulletin 2172–G

**U.S. Department of the Interior
U.S. Geological Survey**

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
Charles G. Groat, Director

U.S. Geological Survey, Reston, Virginia: 2003

This report is only available on-line at:
<http://pubs.usgs.gov/bul/b2172-g/>

For more information about the USGS and its products:
Telephone: 1-888-ASK-USGS
World Wide Web: <http://www.usgs.gov/>

Any use of trade, product, or firm names in this publication is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, it contains copyrighted materials that are noted in the text. Permission to reproduce those items must be secured from the individual copyright owners.

Contents

Abstract	1
Introduction	1
Acknowledgments	1
Concepts and Nomenclature	1
Data	3
Estimation of Inferred Reserves: Statistical Approach	3
Bootstrap Procedures	4
Approximate Nonparametric Confidence Intervals	4
Summary and Implications	6
References Cited	7
Appendix	8

Figures

1. Diagram showing layout of a discovery table for gas in gas fields.....	2
2. Sketch showing wells leading to additions to reserves in discovered fields.....	2
3. Graph showing cumulative monotone field growth functions, from 1995 U.S. Geological Survey assessment	3
4. Graphs showing base case annual additions to oil and gas reserves from oil and gas fields discovered before 1992	5
5. Histograms showing bootstrap distribution of inferred oil reserves and inferred gas reserves for the 80-year period from 1992 through 2071, for fields discovered prior to 1992	6

Table

1. Estimated inferred reserves, with 95 and 5 percent confidence bounds from fields discovered as of December 31, 1991, onshore and State water areas of the conterminous United States	7
---	---

Uncertainty and Inferred Reserve Estimates— The 1995 National Assessment

By Emil D. Attanasi and Timothy C. Coburn¹

Abstract

Inferred reserves are expected additions to proved reserves of oil and gas fields discovered as of a certain date. Inferred reserves accounted for 65 percent of the total oil and 34 percent of the total gas assessed in the U.S. Geological Survey's 1995 National Assessment of oil and gas in onshore and State offshore areas. The assessment predicted that over the 80-year period from 1992 through 2071, the sizes of pre-1992 discoveries in the lower 48 onshore and State offshore areas will increase by 48 billion barrels of oil (BBO) and 313 trillion cubic feet of wet gas (TCF). At that time, only point estimates were reported. This study presents a scheme to compute confidence intervals for these estimates. The recentered 90 percent confidence interval for the estimated inferred oil of 48 BBO is 25 BBO and 82 BBO. Similarly, the endpoints of the confidence interval about inferred reserve estimate of 313 TCF are 227 TCF and 439 TCF. The range of the estimates provides a basis for development of scenarios for projecting reserve additions and ultimately oil and gas production, information important to energy policy analysis.

Introduction

Inferred reserves are expected future additions to proved reserves in identified fields. These resources accounted for 65 percent of the total oil and 34 percent of the total gas assessed in the U.S. Geological Survey's (USGS) 1995 National Assessment of oil and gas resources in United States onshore and State offshore areas (Gautier and others, 1996). The assessment predicted that during the 80-year period from 1992 through 2071, sizes of pre-1992 conventional oil and gas discoveries in the onshore and State offshore areas of the conterminous United States *would increase* by 48 billion barrels of crude oil (BBO) and 313 trillion cubic feet of wet gas (TCF) (290 TCF dry gas). Inferred reserve estimates were calculated by the extrapolation of historical trends rather than the application of geologic concepts. Because the computations to derive statistical estimates of inferred reserves are complex, only point estimates

were reported and the uncertainty of those estimates was not addressed.

In recent years inferred reserves have been the dominant source of the Nation's new proved reserves. Proved reserve levels are important because they are leading indicators of the sustainability of production. Because no more than 10 to 15 percent of the proved reserves in conventional fields can be extracted annually without risking reservoir damage, proved reserves limit annual production to an amount well below recoverable resource volumes. Accordingly, production forecasts should also reflect the uncertainty associated with statistically based inferred reserve estimates.

This report presents uncertainty estimates for the 1995 inferred reserves values. The method to obtain the uncertainty estimates represents an application of the bootstrap theory (Efron and Tibshirani, 1998) and the algorithm used to calculate inferred reserves for the 1995 USGS Assessment (Root and others, 1997). The report is organized as follows: (1) data and method used for estimating inferred reserves; (2) the bootstrap procedures applied to compute the confidence intervals; (3) the implications of uncertainty in inferred reserve estimates to energy policy analysis.

Acknowledgments

The authors thank Dave Menzie, David Root, John Schuenemeyer, and Ione Taylor for their reviews of the methodology and the report. Lorna Carter's editorial comments helped to clarify the report.

Concepts and Nomenclature

Inferred oil and gas reserves are defined here as cumulative future additions to proved reserves in oil and gas fields discovered as of a certain date. *Proved reserves* are estimated quantities of hydrocarbons that geologic and engineering data demonstrate with reasonable certainty to be recoverable from identified fields under existing economic and operating conditions (Society of Petroleum Engineers, 1987; Securities and Exchange Commission, 1981). Inferred reserve estimates may specify the time period during which the additions are expected to occur. Here, "*field growth*" refers to annual additions to

¹Abilene, Texas 79699-9325.

2 Geologic, Engineering, and Assessment Studies of Reserve Growth

reserves in previously identified fields. *Known recovery*, the estimate of field size used here, is defined as the sum of a field's past production and its current estimate of proved reserves.

Analysts have frequently observed that when fields are grouped by year of their discovery (or vintage), the sum of the field size estimates (known recoveries) for each vintage tends to increase systematically over time. Figure 1 shows a partial discovery table that arrays the sum of the estimates of known recoveries of gas in gas fields by the particular vintage year shown in the first column. For each row or vintage year, entries from left to right represent the sum of estimates of field

sizes (known recovery) as of 1977, 1978, ... to the last year of data. The systematic change across estimate years suggests that a statistical approach to modeling the process of field growth might be appropriate.

During field development, proved reserves increase, and hence estimates of known recovery increase, as delineation drilling extends the boundary of the productive area. Estimated field size increases as the lateral boundary of the field is extended and as newly discovered pools that are shallower or deeper are added (fig. 2). Estimated field size also increases with improved recovery of in-place hydrocarbons resulting from well stimulation and fluid injection.

Discovery year	1977	1978	1979	1980	1981	1982	No. fields by vintage
1975	2,538	2,873	2,944	3,196	3,260	3,303	317
1976	1,555	1,714	1,841	2,156	2,417	2,505	340
1977	1,323	2,087	2,630	2,881	3,285	3,617	509
1978	0	1,012	1,801	2,752	2,952	3,500	422
1979	0	0	922	1,791	2,799	3,067	438
1980	0	0	0	784	1,533	2,252	453
1981	0	0	0	0	891	1,441	491

Figure 1. The layout of a discovery table: each row represents the sum of field sizes for each vintage (all fields with the same discovery year). Entries from left to right are successive estimates of total volume of gas in all gas fields in each vintage. Far right column shows the number of non-associated gas fields for each vintage. Source of data is the 1991 Oil and Gas Integrated Field File issued in 1993 for fields discovered prior to 1992. (Units are in billions of cubic feet.)

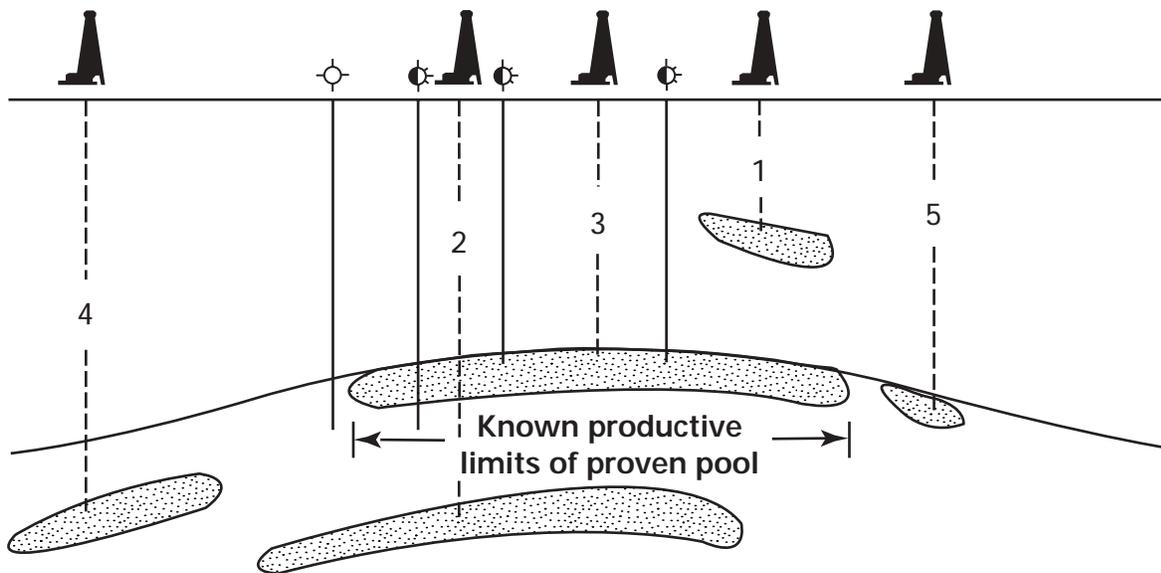


Figure 2. Schematic of wells leading to additions to reserves in discovered fields. 1, shallower pool test; 2, deeper pool test; 3, infill well; 4, new pool test; 5, extension or outpost (modified from Drew, 1997). In practice, the operator or regulatory body may classify accumulations penetrated by wells 1 through 5 as a single field or as more than one field. Recognition of relationships among accumulations could also be further complicated by order in which wells were actually drilled. Open symbol, dry hole; half-filled symbol, successful producers.

The U.S. Department of Energy defines a field as an area consisting of a single or multiple reservoirs all grouped or related to the same individual geologic structure and (or) stratigraphic feature. Two or more reservoirs that occur in the field may be separated by impermeable strata vertically, laterally, or both (Energy Information Administration, 1998). In practice, the definition of an oil or gas field is not exact. Pools are not always assigned to fields for well-defined geologic reasons; assignments may be for regulatory convenience or may reflect the discovery sequence of the pools. Depending on the sequence of the discovery of pools, accumulations shown in figure 2 may be assigned to a single field or to more than one field.

For any vintage, the *annual growth factor* for fields of age i is the ratio of the aggregate estimated volume of known recovery at $(i+1)$ years after discovery to the aggregate estimated volume of known recovery i years after discovery. The *cumulative growth factor* is the ratio of the aggregate estimates of field sizes m years after discovery to the aggregate initial estimates of known recovery. Cumulative growth factors can also be computed as the successive products of annual growth factors. For each vintage, the cumulative growth function is represented by the series of cumulative growth factors computed for each successive year after discovery (fig. 3). The appendix shows the mathematical expressions of the growth function and the procedure used to compute it.

Data

The inferred reserve estimates of the U.S. Geological Survey's 1995 National Assessment were based on estimates of known recovery of conventional oil and gas fields discovered before 1992 and located in onshore areas and State waters of the United States (fields in the Federal Outer Continental Shelf are excluded). The basic data are the Oil and Gas Integrated Field File (OGIFF). The U.S. Energy Information

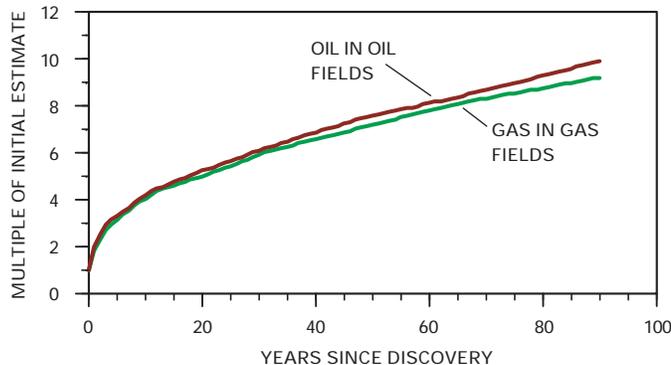


Figure 3. Cumulative monotone field growth functions for oil in oil fields and gas in non-associated gas fields. Functions are determined using data from U.S. Energy Information Administration's Oil and Gas Integrated Field File (issued in 1993 with data through 1991). (See Root and others, 1997.)

Administration (EIA) compiles these data from proprietary industry responses to the "Annual Survey of Domestic Oil and Gas Reserves" (EIA-Form 23). The compiled data are a series of annual estimates, starting in 1977, of individual field sizes in volumes of both crude oil and wet gas. The EIA data consist of short series of estimates of field sizes rather than complete histories of fields from discovery to abandonment.

Fields are classified as non-associated gas if the ratio of the estimated gas recovery to oil recovery is at least 20,000 cubic feet of wet gas to one barrel of oil. Otherwise a field is assumed to be an oil field. Fields discovered before 1901 and those without discovery dates are excluded. Field age is the number of years since discovery. Production begins for most onshore lower 48 fields soon after discovery.

Estimation of Inferred Reserves: Statistical Approach

The 1995 procedure for estimation of inferred reserves started with the preparation of a discovery table (fig. 1) showing estimated volumes of gas in conventional² gas fields by vintage and year of estimate. A similar table was constructed for oil in conventional oil fields. Cumulative growth functions (fig. 3) were then determined using the 1993 version of the OGIFF (field estimates were as of the end of 1991). The method for determining the cumulative growth function amounted to a constrained optimization procedure described in detail in Attanasi and Root (1994). Two assumptions were made: on average, (1) the total known recovery of fields in specific vintage does not shrink as with age, and (2) the total known recovery in older vintages cannot grow by a larger factor in 1 year than total known recovery of a younger vintage. With these assumptions imposed as constraints, cumulative growth functions will be monotone (Appendix). After each vintage has reached a maximum of 91 years³ since discovery, it is no longer assumed to grow.

In summary, the approach to estimation of inferred reserves first requires that cumulative growth factors be calibrated and then used with the latest year estimate of known recovery for fields in each vintage to generate annual additions to reserves through the next 80 years (1992 to 2071). Specifically, growth factors and the 1991 estimates⁴ of known recovery of oil and gas fields were used to compute annual additions to reserves through an 80-year period starting in 1992. The sum of annual estimates from 1992 through 2071 constitutes the 80-year point estimate of inferred reserves.

²OGIFF data have no geologic information associated with them. Fields producing principally from unconventional accumulations were identified by geologists and excluded from the data used to determine the cumulative growth functions and to estimate inferred reserves.

³The 91 years represents the time span over which the growth factors were computed (1901–1991).

⁴At the time of the 1995 assessment, the 1991 estimates were the latest estimates of field sizes in OGIFF.

The series of annual reserve additions and the inferred reserve estimates at 30 and 80 years are referred to here as the *base case*. Annual additions to proved reserves of (associated) gas from discovered oil fields were obtained by applying the cumulative growth function for oil to the known recovery of associated gas.⁵

Bootstrap Procedures

The bootstrap is a method for constructing the sampling probability distribution of sample statistic when the probability distribution governing the population is unknown. Here, the statistic of interest is represented by the *base case* estimate of inferred reserves (either oil or gas) for the 80-year period. The sampling distribution is ultimately used to construct confidence limits for the sample statistic (that is, the *base case* estimate). The motivation for the application of the bootstrap is the following.

Suppose we want to determine the sampling distribution of a statistic, $T(\mathbf{x})$, which is calculated from observed data, \mathbf{x} a vector. If $T(\mathbf{x})$ is a simple analytic function of the data and the probability distribution $F_{\mathbf{x}}$ that generates \mathbf{x} is known, a theoretical approach could be used to obtain an exact or approximate expression of the sampling distribution of $T(\mathbf{x})$. If $T(\mathbf{x})$ is a complicated function of the observations (as is the *base case* estimate of inferred reserves) but the probability distribution generating the observations is known, standard Monte Carlo simulation could be used to numerically generate the probability distribution for $T(\mathbf{x})$. But for this case, the probability distribution $F_{\mathbf{x}}$ is unknown.

The bootstrap method is based on the idea of “resampling with replacement” the original sample, F^* , of n observations on which the statistic of interest, $T(\mathbf{x})$, is computed. Data values are assumed to be equally likely events. For each bootstrap replicate the statistic $T(\mathbf{x})$ is computed, and is now denoted $T(x)$ to indicate its association with the bootstrap replicate. The procedural requirement to sample with replacement leads to the possibility of generating all possible combinations of data (Efron and Tibshirani, 1998). If the original sample data are such that F^* (the original empirical distribution) reasonably represents $F_{\mathbf{x}}$ (the population), then with a sufficient number of bootstrap replications, the sampling distribution of $T(x)$ will approximate the actual sampling distribution of $T(\mathbf{x})$ reasonably well.

The bootstrap procedure was applied as follows. The original sample is represented by the data used in the base case. A stratification scheme was devised so that each vintage (set of fields with a common discovery year) represented in the original discovery table comprised one of the strata. Field records (the series of 15 estimates, 1977 through 1991, of individual field sizes) were grouped by vintage, and the records for each vintage were sampled with replacement to construct a bootstrap replicate of size n_i , where n_i is the original (total) number of field records for that vintage (last column of fig. 1)⁶. The field records contained in each bootstrap replicate

were summed by vintage and used to construct a bootstrap discovery table. An individual discovery table was constructed for each such bootstrap sample.

To obtain the j^{th} bootstrap estimate of inferred reserves, $T_j(\mathbf{x})$, the monotone cumulative growth function was determined using the associated bootstrap discovery table, and annual field growth was calculated for an 80-year period from 1992 through 2071. To assure stability of the confidence intervals, 2,000 bootstrap replicate discovery tables, growth functions, and field growth calculations were computed for both oil and gas. The procedure was applied to oil and gas fields separately. The 2,000 bootstrap replicates for estimates, $T_j(\mathbf{x})$, of inferred reserves for oil constituted the bootstrap distribution for oil in oil fields for the 80-year period. Similarly, the 2,000 bootstrap replicates of the estimates of inferred gas formed the bootstrap distribution for gas in gas fields for the 80-year period. The bootstrap distributions of $T(\mathbf{x})$ were the basis for construction of approximate 90 percent confidence intervals for the *base case* estimates of the 80-year estimates of inferred reserves. The 30-year inferred reserve estimates were also computed.

For oil and non-associated gas, respectively, figure 4A and figure 4B show the *base case* annual field growth calculations, along with the mean and median of all 2,000 bootstrap replicates. Parts A and B of figure 5 are the bootstrap distributions of 80-year growth (cumulative additions to reserves) for oil and total gas, respectively. The additions to proved reserves of (associated) gas from discovered oil fields are calculated by application of the cumulative growth function for oil to the known recovery of associated gas.⁷ The distributions representing 80-year inferred reserves of associated gas and non-associated gas were combined to compute the total gas distribution (fig. 5B) using Monte Carlo simulation.

Approximate Nonparametric Confidence Intervals

The distributions of bootstrap estimates of inferred reserves were used to compute approximate confidence intervals for the *base case* estimates. Because the forms of the bootstrap distributions as shown by the histograms in figure 5A and B, were not readily identified with any known parametric distribution, a nonparametric approach to confidence

⁵To preserve the distinction between the sources of additions to reserves, that is, oil or gas fields, the function used to calculate reserve additions of oil in oil fields was used to calculate associated gas reserve additions in oil fields, and the function used to calculate gas reserve additions in gas fields was also used to compute natural gas liquid reserve additions in gas fields.

⁶Each vintage may be visualized as an urn containing the balls corresponding to the n_i original field records of that vintage. The sampling procedure was to sample each urn with replacement n_i times.

⁷Extremely small amounts of oil in gas fields were reported in the OGIFF, and the changes in known recovery for this byproduct were assumed to follow the cumulative growth function for non-associated gas. The totals shown are already aggregated (probabilistically) in a fashion similar to that of the aggregation of non-associated and associated gas.

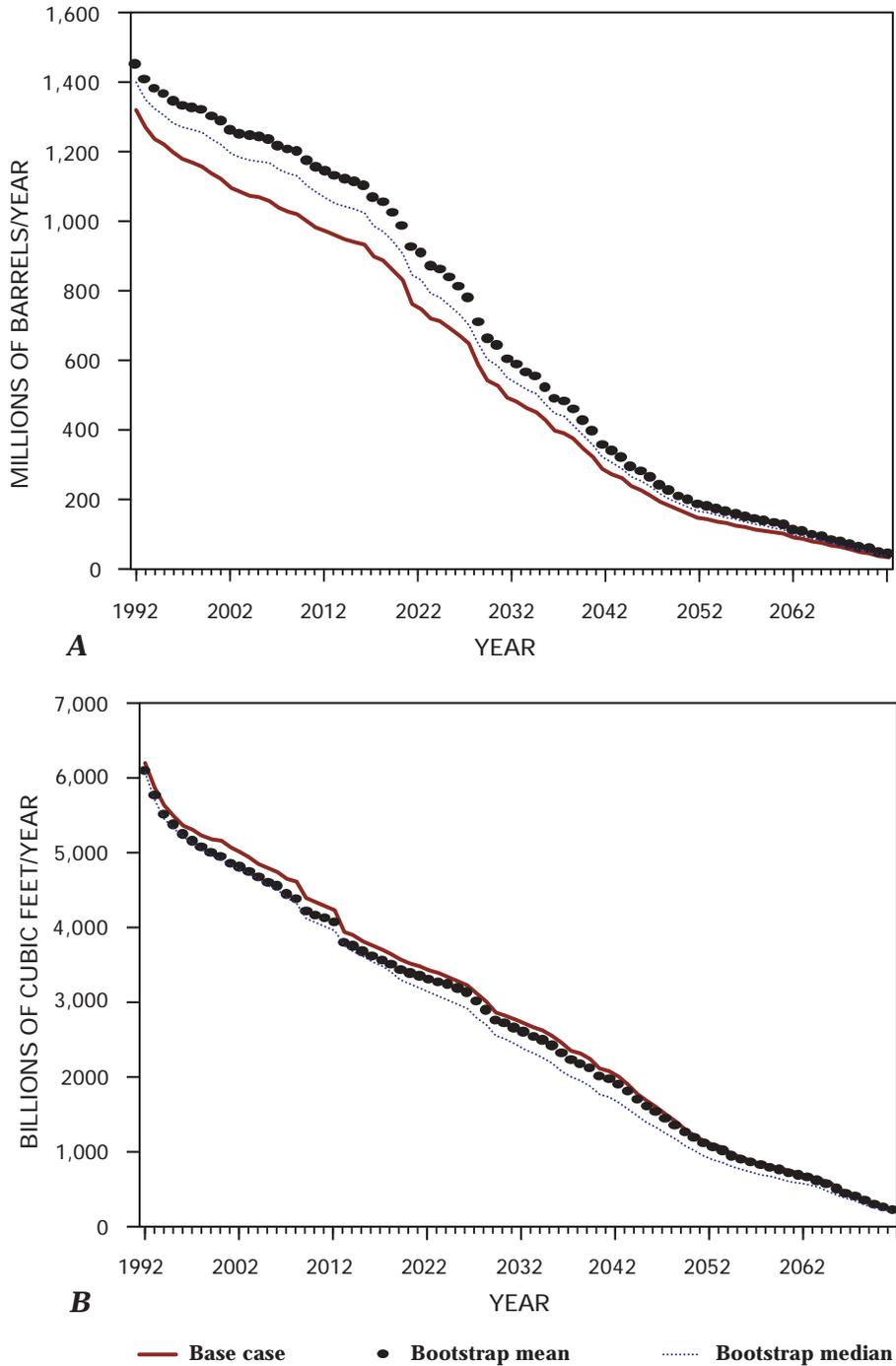


Figure 4. Base case annual estimates of additions to *A*, oil reserves from oil fields, and *B*, gas reserves from non-associated gas fields, discovered before 1992 in lower 48 onshore and State water areas, along with the mean and median of the associated bootstrap distribution.

interval construction was employed. The medians of the bootstrap distributions shown in figure 5A and B deviate slightly from the *base case*⁸; consequently the approximate confidence intervals based on the bootstrap distributions were recentered about the *base case*. To accomplish this, the bias-correction procedure for the percentile method of confidence interval

construction outlined by Efron and Gong (1983) was applied. The method proceeded as follows.

The *base case* inferred reserve estimate in year k (30-year 2021 or 80-year 2071) is defined to be \hat{p}_k . If the bootstrap distribution is centered at \hat{p}_k , the endpoints of a $1-2\alpha$ central confidence interval on \hat{p}_k are just the α and $1-\alpha$ percentiles of

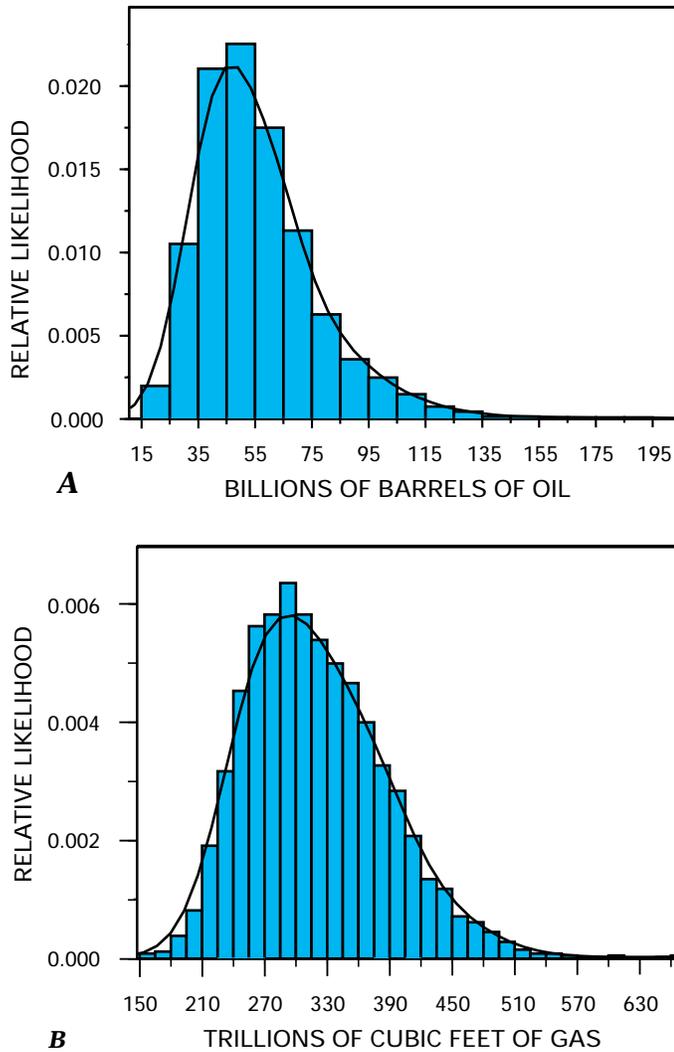


Figure 5. Bootstrap distribution of **A**, cumulative additions to oil reserves from fields discovered before 1992 in lower 48 onshore and State water areas for the 80-year period from 1992 through 2071; **B**, cumulative additions to gas reserves from oil and gas fields discovered before 1992 in lower 48 onshore and State water areas for the same 80-year period as **A**. Smooth curves plot the implied density functions.

that distribution. When the distribution is not centered on \hat{p}_k , as in this case, a (bias) correction must be applied to the endpoints of the interval to ensure probability coverage of $1-2\alpha$.

The adjustment is based on standardizing the difference between \hat{p}_k and the median of the empirical bootstrap distribution in the following manner. Suppose $Q(\hat{p}_k)$ is the proportion of the empirical distribution that is smaller in value than \hat{p}_k . Let

$$z_0 = S^{-1}(Q(\hat{p}_k)) \quad (5)$$

where S^{-1} is the inverse function of the standard normal distribution.

The bias-corrected $1 - 2\alpha$ central confidence interval on \hat{p}_k is given by

$$[Q^{-1}(S(2z_0 - z_\alpha)), Q^{-1}(S(2z_0 + z_\alpha))], \quad (6)$$

where z_α is the $100(1-\alpha)$ percentile of the standard normal distribution (Efron and Gong, 1983), S is the standard normal cumulative distribution, and Q^{-1} is the inverse of the bootstrap distribution. If the bootstrap distribution is centered on \hat{p}_k , then $Q(\hat{p}_k) = 0.5$ and z_0 in equation (5) equals zero. The recentering of the distribution becomes more significant as the magnitude of the difference between \hat{p}_k and the median of the bootstrap distribution increases.

Table 1 shows *base case* estimates of the cumulative additions to oil and gas reserves for 30 (1992–2021) and 80 (1992–2071) years of growth (Root and others, 1997), along with endpoints of the associated recentered 90 percent confidence intervals. To compute the *base case* estimate of total inferred gas, the *base case* point estimates of non-associated and associated gas were simply added. However, in order to obtain proper confidence intervals, the bootstrap distribution of the non-associated gas inferred reserve estimates and distribution of the associated gas inferred reserve estimates were combined using Monte Carlo simulation. The bootstrap distributions for non-associated gas and associated gas were derived independently, and growth functions for gas fields and oil fields were determined separately. In the absence of strong evidence to the contrary, the aggregation assumed statistical independence of the non-associated gas and associated gas distributions.

According to table 1 for the *base case*, the 30-year inferred oil reserve estimate represents 69 percent of the corresponding 80-year calculation. The 30-year calculation of the total gas inferred reserve estimate is 63 percent of the corresponding 80-year estimate. This is consistent with the profiles of additions to reserves shown in figure 4A and B. For the 80-year estimates, endpoints for the recentered confidence interval for *total oil* are 25 BBO and 82 BBO, respectively, with the *base case* point estimate at 47 BBO. The non-associated gas recentered confidence interval endpoints are 153 TCF and 356 TCF, respectively, and the *base case* point estimate is 223 TCF. The endpoints of the recentered 90 percent confidence interval for *total gas*, over the 80-year period, are 227 and 439 TCF, respectively and the *base case* estimate is 313 TCF (293 dry gas). The results presented in table 1 and the distributions shown in figure 5 convey more information about the properties of the inferred reserve estimates than the singular point estimates reported in the past.

Summary and Implications

Note that the specific focus of this report is to provide a method for quantifying uncertainty that is unique to the problem of estimating inferred reserves. The estimation of inferred reserves departs from conventional parametric statistical methods because of the complexity of the problem and the inability to obtain a solution using traditional analytic methods. The

⁸For total oil the *base case* and median are 48.1 BBO, and 52.0 BBO, and the total gas *base case* and median values are 313 and 312 TCF, respectively.

Table 1. Estimated inferred reserves, with 95 and 5 percent confidence bounds from fields discovered as of December 31, 1991, onshore and State water areas of the conterminous United States.

	30-year growth		
	5%	<i>base</i>	95%
Total oil (billions of barrels)	19.1	33.1	52.1
Associated gas (trillions of cubic feet)	29.7	59.2	102.9
Non-associated gas (trillions of cubic feet)	102.0	138.6	193.0
Total gas (trillions of cubic feet)	151.3	197.9	254.9

	80-year growth		
	5%	<i>base</i>	95%
Total oil (billions of barrels)	25.3	48.1	82.3
Associated gas (trillions of cubic feet)	40.3	90.4	174.5
Non-associated gas (trillions of cubic feet)	152.6	222.6	356.1
Total gas (trillions of cubic feet)	227.1	313.0	439.4

inferred estimation procedure initially relies on a form of constrained optimization to determine the cumulative growth function, where the basis of the optimization is the minimization of the sum of squared differences between estimated and actual values within the data. No assumptions are made about the underlying distribution of the differences, the presence or absence of correlation among them, or what they physically represent. The bootstrap procedure is used to obtain numerical, rather than analytical, characterization of the distribution of inferred reserve estimates. Finally, the bootstrap distributions of inferred reserve estimates provide the basis for constructing approximate confidence intervals that convey uncertainty about the single point estimate; and because the form of the bootstrap distributions cannot be established analytically, a nonparametric approach to confidence interval construction was employed.

The individual field records spanned a 15-year period (1977 through 1991), which include times of boom and bust “typical” of the petroleum industry. The comparison of short-term predictions with any particular near-term realization is not likely to be favorable because short-term proved reserve additions are dominated by immediate economic conditions. During a period of 15 years or longer, the dominant part of the changes in field size is related to the extension and discovery of new pools and the application of improved technology.

The range of the estimates provides a rational basis for development of scenarios for projecting reserve additions and ultimately oil and gas production. Well-thought-out scenarios are often more appropriate than standard sensitivity analysis because these resource markets tend to react to change in an erratic fashion. Future applications of the bootstrap to the inferred reserve estimation problem could focus on the question of how the quality of the estimate changes as the amount of data, that is, years of discovery and number of years of estimation, increases. The uncertainty intervals could also be used as one of several criteria for evaluating alternative methods for estimating inferred reserves.

References Cited

- Attanasi, E.D., and Root, D.H., 1994, The enigma of oil and gas field growth: American Association of Petroleum Geologists Bulletin, v. 78, no. 3, p. 321–332.
- Drew, L.J., 1997, Undiscovered petroleum and mineral resources—Assessment and controversy: New York, Plenum, 210 p.
- Efron, Bradley, and Gong, Gail, 1983, A leisurely look at the bootstrap, the jackknife and cross-validation: The American Statistician, February 1983, v. 7, no. 1, p. 36–48.
- Efron, Bradley, and Tibshirani, Robert, 1998, An introduction to the bootstrap: Boca Raton, Fla., Chapman and Hall/CRC, 436 p.
- Energy Information Administration, 1998, U.S. crude oil, natural gas, and natural gas liquids reserves; 1997 Annual report: DOE/EIA-0216(97), 158 p.
- Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1996, 1995 National assessment of United States oil and gas resources—Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series 30, version 2 corrected.1 disk.
- Root, D.H., Attanasi, E.D., Mast, R.F., and Gautier, D.L., 1997, Estimates of inferred reserves for the 1995 U.S. Geological Survey National Oil and Gas Resource Assessment: U.S. Geological Survey Open-File Report 95–75–L, 29 p.
- Securities and Exchange Commission, 1981, Regulation S-X Rule 40-10, Financial Accounting and Reporting Oil and Gas Producing Activities, Securities and Exchange Commission Reserves Definitions, March 1981: New York, Bowne and Company Inc.
- Society of Petroleum Engineers, 1987, Reserves definitions approved: Journal of Petroleum Technology, v. 39, no. 5, p. 576–578.

Appendix

For each vintage, the cumulative growth *function* is represented by the series of cumulative growth *factors* computed for each successive year after discovery. Changes in estimated known recovery conform to the relation

$$\hat{c}(i, j+k, k) = c(i, j) [G(j+k-i)/G(j-i)] + d(i, j, k), \quad (1)$$

where $c(i, j)$ is the aggregate known recovery in year j of vintage i fields (that is, fields discovered in year i), k is elapsed time (in years where $k > 0$), $\hat{c}(i, j+k, k)$ is the estimate of aggregate known recovery in year $j+k$ of vintage i fields, $d(i, j, k)$ is a deviation (or error term), and $G(\cdot)$ is the cumulative growth factor.

Relation (1) implies that, for a specified vintage with discovery year i , the aggregate of the known field recoveries in year $j+k$ can be determined by multiplying the aggregate of the estimated field recoveries in year j by an appropriate factor that is the ratio of cumulative growth factors for field ages $(j+k-i)$ and $(j-i)$. $G(n)$ is bounded, that is $G(0) = 1$ and $G(n)$ is

a constant for n greater than 91, which is the length of the data series (1901–91).

The cumulative growth factors, $G(j+k-i)$, are computed using an optimization algorithm that minimizes the sum of squared deviations,

$$ssd = \sum d(i, j, k)^2, \quad (2)$$

between all combinations of the actual and estimated aggregate field recoveries, $\hat{c}(d, j+k)$, within the data (that is, discovery table) used in the analysis, where

$$d(i, j, k) = c(i, j+k, k) - \hat{c}(i, j+k, k). \quad (3)$$

The optimization is described in Attanasi and Root (1994). The best growth function is the set of cumulative growth factors that minimizes ssd .

The following condition is also imposed on the cumulative growth coefficients,

$$1 \leq G(m+1)/G(m) \leq G(m)/G(m-1), \quad m \geq 1. \quad (4)$$

The left side of the inequality ensures that vintages do not shrink as they age. The right side of the inequality implies that an older vintage cannot grow by a larger factor in one year than a younger vintage.